

UM ESTUDO SOBRE A IMPORTÂNCIA DA TECNOLOGIA BIG DATA E PERSPECTIVAS FUTURAS DE SUA APLICAÇÃO NA INDÚSTRIA.

Jhonatan Kayo Silva Sena ¹, Paulo Henrique Lopes Silva ²

Resumo: Vivemos o que é atualmente chamado por estudiosos de Indústria 4.0, a quarta revolução industrial, fortemente ligada a automação de processos com o auxílio da análise de dados dos mais diversos equipamentos conectados. O termo Big Data é utilizado na literatura especializada para referir-se a esse enorme volume de dados armazenados em constante crescimento. O presente artigo, de caráter exploratório, visa explicar inicialmente o que é Big Data e posteriormente discutir os impactos que a aplicação do método Big Data tem apresentado na indústria e mercado. Quais seriam os custos e ferramentas necessárias, se o mercado está preparado para aplicá-lo nos meios de produção e distribuição de serviços e o que aguardar para o futuro da sociedade com a evolução dessa tecnologia aplicada a um novo formato de indústria, fortemente ligada a internet das coisas e a produção baseada em decisões multicritério.

Palavras-chave: Big Data; Indústria 4.0; Internet das Coisas; Automação.

Abstract: We live what is currently called by researchers of industry 4.0, the fourth industrial revolution, strongly linked to process automation with the aid of data analysis of the most diverse connected equipment. The term Big Data is used in the literature to refer to this huge amount of stored data that constantly grows. This exploratory article aims to explain what Big Data is, and after that to discuss the impacts that the application of the Big Data method has done on industry and the market. What would be the costs and tools needed, if the market is prepared to apply it in the means of production and distribution of services and what to look forward for the future of society with the evolution of this technology applied to a new format of industry, strongly linked to the internet of things and production based on multi criteria decisions.

Keywords: Big Data; Industry 4.0; Internet of Things; Automation.

1. INTRODUÇÃO

Na primeira década do século XXI, vivenciamos um *boom* tecnológico que nos levou a miniaturização destes equipamentos, como *smartphones* e computadores, que ficaram cada vez mais portáteis e que aliados a linguagens de programação cada vez mais sofisticadas nos trouxeram o advento de itens não só pequenos mas também robustos tecnologicamente, possibilitando uma comunicação e troca de dados ágil e em qualquer localização por meio da internet.

Em um mundo cada vez mais conectado, a quantidade de dados a se lidar tem crescido exponencialmente, numa velocidade exponencial [1]. Nesta década, com a industrialização atingindo patamares cada vez mais sofisticados e mais que nunca, focada na maximização da produção de forma inteligente, os processos produtivos voltam-se para uma produção mais dinâmica e diversificada. Muitas vezes as indústrias têm adotado um modelo de customização em massa e a análise combinatória e inteligente de dados por meio de softwares aliados a ferramentas de controle e captação de dados tem ganhado seu espaço.

Essa indústria mais dinâmica, autônoma e que usa de forma robusta dados a fim de tornar a produção mais assertiva recebeu o nome de indústria 4.0 pelo governo alemão em 2011 [6].

O impacto da Indústria 4.0 vai muito além da digitalização de informações, ela é uma forma complexa de inovação que se baseia na combinação de múltiplas tecnologias, que modificam severamente o processo dentro de uma organização e até mesmo na hierarquia da mesma [2]. É possível apontar, por exemplo, que o processo de padronização da produção que vivenciamos na última revolução industrial, agora se volta para um processo padronizado e com alta capacidade lógica, graças a sua facilidade na manipulação de dados, mas que não é centralizado a uma linha padronizada de produção em escala, mas voltada a um modelo enxuto de produção e

1. Acadêmico do curso de bacharelado em ciência e tecnologia da Universidade Federal Rural do Semiárido, e-mail: kayojhonatan@gmail.com

2. Professor Doutor do Departamento de Computação da Universidade Federal Rural do Semiárido, email: phenrique@ufersa.edu.br

que pode adaptar-se aos padrões do mercado sem grandes dificuldades após sua total implementação.

Nos campos da ciência, onde se pode lidar com uma carga de dados tão extensa e variada, vinda das mais diversas fontes, é importante entendermos como o advento desta tecnologia tem impactado na sociedade. Este impacto pode ser sentido no âmbito popular com a popularização da IoT, termo empregado ao grande volume de dispositivos conectados na segunda década deste século, e também no âmbito acadêmico, que conforme estudos, cada vez mais tem se dedicado a evoluir na área de qualidade de dados [3]. Então, tendo-se ciência da multidimensionalidade que a qualidade de dados oferece em termos de aplicação, precisamos compreender como em um mundo totalmente conectado e em que tudo torna-se informação, tecnologias que usam da Big Data seriam a solução para a realização de análises apuradas de conjuntos e estruturas de dados, que outrora eram metódicos e lineares e se tornaram extensos e complexos. Sabendo disso, é necessário entender como esses métodos são aplicados e quais as dificuldades na aplicação e perspectivas futuras sobre o tema.

Diante do que foi exposto, esse estudo tem como objetivo estudar a produção acadêmica sobre o assunto Big Data, e avaliar os impactos que a aplicação da mesma tem na indústria 4.0 e o que seria necessário para aplicação de ferramentas baseadas no método, desde o capital humano a infraestrutura, estabelecendo sua importância, e com isso esquematizar perspectivas futuras sobre o tema. Além disso, quantificar, de acordo com as áreas que a Indústria 4.0 proporciona, quais delas mais tem atuado em conjunto com a análise de dados, apresentando também tendências para o futuro.

2 METODOLOGIA

A metodologia do presente artigo, devido a suas características em relação à análise e explanação dos dados, se dará como uma pesquisa do tipo *exploratória*. Para que fosse possível embasar as informações necessárias para a evolução do tema foi feito um estudo bibliográfico da literatura disponível em meios eletrônicos, fazendo um levantamento das informações publicadas e chegando a uma conclusão sobre o material disponível até o momento [12]. A seleção desses artigos foi feita por meio da ferramenta de busca disponibilizada pelos sites de consulta de acervos acadêmicos, sendo o único critério para seleção dos artigos tratar de indústria 4.0 e Big Data. Para Silva, essa modalidade de pesquisa consiste na revisão de material já publicado, desde livros a periódicos acadêmicos dos mais diversos, porém, pelo caráter que a pesquisa apresenta, visando discutir um tema que vem se desenvolvendo na sociedade acadêmica e de grande aplicabilidade na indústria e comércio, o foco será dado a artigos do campo acadêmico.

Quanto a sua natureza, trata-se de uma pesquisa básica, de caráter qualitativo quanto a sua abordagem. Isso porque o presente material apresenta informações que podem ser aplicadas a um projeto que envolva o conteúdo nela contido, porém objetivando debater o conhecimento compartilhado por meio da comunidade acadêmica a respeito do tema [12]. O material é revisado e encaixando-se no tema proposto, é inserido e problematizado no campo de impactos que a aplicação dos métodos tiveram na área abordada pelo mesmo, seguindo alguns critérios, que serão descritos adiante para trazer uma análise coesa e precisa em relação ao que a pesquisa acadêmica se propõe.

A análise da literatura compreenderá artigos do período de 2011 a 2019, isso porque, como apontado anteriormente, a abordagem do termo “indústria 4.0” iniciou em 2011 [6]. Serão avaliados os artigos em que houve a aplicação da problemática em questão. Para identificação da literatura aplicada, foi feito o uso de páginas que disponibilizam conteúdo acadêmico para leitura, utilizando-se da base de periódicos da CAPES¹, que dá acesso a conteúdo acadêmico de diversas bases do mundo todo e artigos da ABEPRO². Como se trata de um assunto em constante evolução, muitos debates são realizados pela comunidade científica e industrial acerca de muitas questões. Neste trabalho são listadas algumas questões que serão usadas como seus instrumentos norteadores. São elas:

- Quais impactos podem ser verificados com a utilização da Big Data na indústria 4.0?
- O que é necessário implementar para que a tecnologia Big Data possa ser utilizada nas organizações?
- Quais as maiores dificuldades para implementação sabendo-se das barreiras enfrentadas?
- A partir dos impactos verificados com a indústria 4.0 e Big Data, quais setores de atividades podem se beneficiar com elas?

3. REFERENCIAL TEÓRICO

1. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

2. Associação Brasileira de Engenharia de Produção

Nesta seção, é abordado o cenário onde a tecnologia Big Data está inserida e é feita uma breve explicação de seu funcionamento, desde a tecnologia necessária em termos de hardware e também de software. Após isso, são abordados artigos que discutem os desafios e uso da tecnologia em cenários diversos, contrastando com as informações teóricas anteriormente abordadas.

3.1 UM OLHAR SOBRE A INDÚSTRIA 4.0 E A IMPORTÂNCIA DA BIG DATA

São três os principais pilares que sustentam a quarta revolução industrial: *Internet of Things* (IoT) ou internet das coisas, *Cyber-Physical Systems* ou sistemas ciber-físicos e a Big Data [2]. Há também seis tecnologias que dão suporte a esses três pilares: Os robôs autônomos, simulação, realidade aumentada, cibersegurança, manufatura aditiva e a nuvem [4]. Todas essas tecnologias são ilustradas na **Figura 1**.

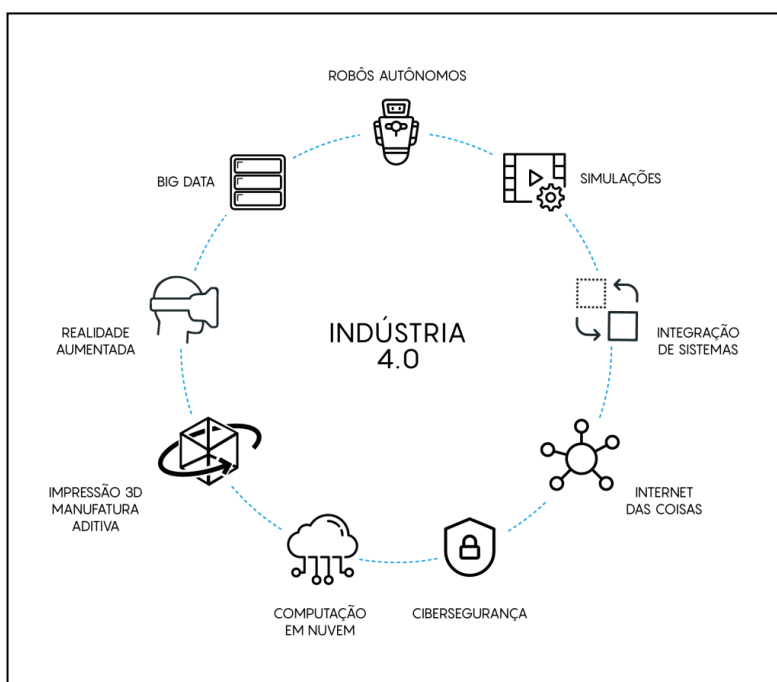


Figura 1. Pilares e tecnologias assistivas da Indústria 4.0; Fonte: xGB Tecnologia Inteligente (2017).

O termo internet das coisas originou-se do que recentemente se identificou como a quantidade significativa de equipamentos conectados à web, até o ano de 2020 é esperado que estes equipamentos conectados estejam na casa dos 50 bilhões, como podemos ver na **Figura 2**. Os sistemas ciber-físicos são sistemas de caráter físico-digital que operam em conjunto entre si, desde as redes de comunicação e nuvem onde estão armazenadas as informações que serão intercambiáveis entre esses sistemas aos minúsculos sensores presentes desde as residências as grandes linhas de produção, permitindo que sistemas que desempenham tarefas de alta complexidade possam ser operados de maneira virtual e algorítmica. Todo esse transitar de informações do campo físico para o digital é tratado por meio do pilar da Big Data, que se trata justamente da análise integrada e qualitativa dos dados coletados e transmitidos a todo o momento [2].

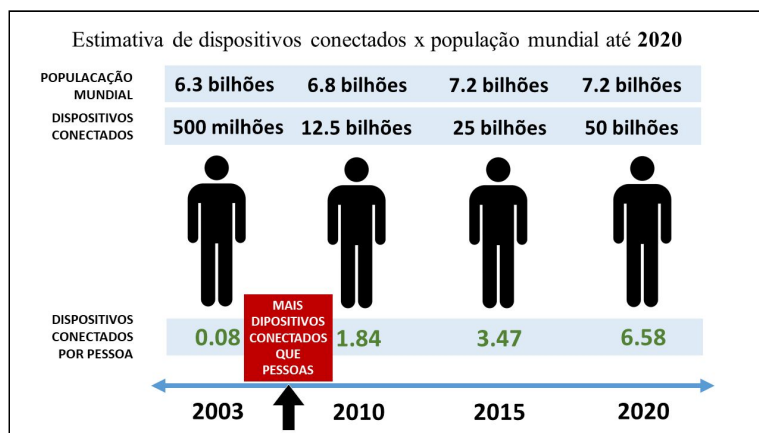


Figura 2. Adaptado de [2] (2018).

A Indústria 4.0 é diretamente relacionada à otimização, digitalização e customização da produção, trazendo consigo automação e adaptação na interação homem máquina e acrescentando valor agregado com serviços e ferramentas de negócios, além de uma transferência automática de dados e comunicação entre dispositivos [5]. O termo “Indústria 4.0” data de 2011, tendo sido mencionado pelo governo alemão como tendência a partir de então para a indústria e vem sendo usado desde então para referenciar a forte tendência de digitalização e automação nos mais diversos âmbitos da sociedade [6].

Foi somente no início da década de 90 que cientistas da computação começaram a considerar o problema de dimensionar, definir e otimizar a qualidade dos dados que eram obtidos e armazenados em bancos de dados (Batini, 2006 apud [7]). Wang e Strong definem em artigo de 1996 (apud [7]) que a qualidade de dados é multidimensional, ou seja, a infinidade de variáveis podem ser utilizadas nos mais diversos âmbitos para as mais diversas finalidades e o que define a qualidade no tratamento desses dados é a necessidade na qual eles serão utilizados. É importante, na qualidade de dados, que os dados retornem informações relevantes e adequadas ao objetivo do processo. Um estudo do processo e a escolha de métodos adequados para o tratamento dos dados é muito importante nessa etapa e as aplicações são diversas. É importante considerar as dimensões da qualidade e possibilidades tecnológicas de aplicação dos dados envolvidos.

Para a indústria e comércio, o surgimento dos métodos de análise e tratamento de dados baseado na Big Data foi muito bem visto, o que explica o crescimento do interesse acadêmico sobre o assunto. O mesmo tornou-se valioso por trabalhar a imensa quantidade de dados que estes setores possuem, maximizando as estratégias de ganho, reduzindo os riscos e mais que isso, fornecendo possibilidades antes inexploradas no campo da inovação, com um contato muito mais próximo com o consumidor final.

No âmbito doméstico e social, estas informações podem ser utilizadas para auxiliar na rotina e objetivos de um indivíduo ou um grupo. O maior trunfo do campo nesta época tem sido a possibilidade de conectar-se a internet presente numa ampla variedade de equipamentos, acessíveis tanto a população em geral quanto às grandes corporações [3]. Até onde a área da ciência de dados irá evoluir é uma incógnita pois as possibilidades crescem junto ao crescimento tecnológico e da nuvem de dados que a internet das coisas tem formado.

No Brasil, conforme ilustrado na **Figura 3**, 58% das indústrias estão cientes da importância da indústria 4.0 para a competitividade que há de vir, onde as empresas que encontram-se em transição para o novo formato de indústria priorizam tecnologias para aumentar a eficiência no processo produtivo e gestão do negócio, porém menos da metade de fato faz uso das tecnologias aplicáveis a ela, o foco tem sido otimizar o processo de produção e aumentar a produtividade [11].

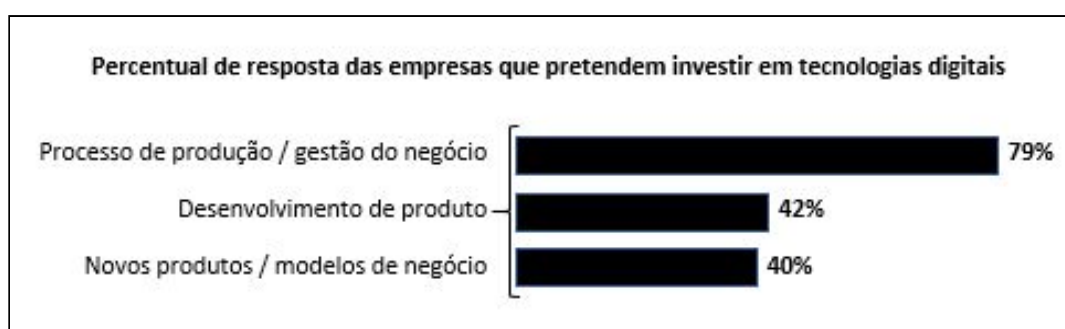


Figura 3. Valores passam de 100% pois respostas podiam ser múltiplas. adaptado de [11] (2018).

3.2 O QUE É BIG DATA

O termo Big Data pode ser definido como um grande amontoado de dados desprovido de estrutura ou organização [16]. Estes dados precisam estar ligados a uma estrutura que possa não só os receber como filtrá-los e direcioná-los de maneira adequada, para isto ser possível, a Big Data segue alguns fundamentos.

Podemos caracterizar a Big Data através de 5 termos, também chamados de 5 V's: volume, variedade, velocidade, veracidade e valor [7]. Nos quais cada um possui um conceito específico e que justifica a sua atribuição quanto ao que é Big Data:

- **Volume:** refere-se ao volume crescente de dados e informações que são gerados das mais diversas fontes. Fagundes faz referência a El, Khatak e Buhler a respeito disso pois estes especificam que o volume de dados gerado pelas soluções baseadas em Big Data não só é substancial, gerando um peso financeiro em relação a estrutura de comunicação e rede, como também crescente, fazendo com que essa estrutura precise ser também escalável [7]. Mais que uma estrutura física, o uso de análise de dados baseada em big data é um desafio também em termos de software, que devem ser ajustados para satisfazer o propósito no qual ele é inserido em meio a uma larga variedade de dados [6].
- **Variedade:** Os dados existem nas mais variadas formas, como texto, áudio, imagem e muitos outros. Essa variedade explica a complexidade de um método robusto de análise de dados [16]. Não se trata mais de apenas um simples banco de dados, o volume variado de dados é obtido não só de registros de máquinas, lógicos e previsíveis, mas também de seres humanos interagindo com sistemas diversos conectados em rede.
- **Velocidade:** Se tratamos de um grande volume de dados em um processo, seja este destinado a produção ou oferta de um serviço, este processo precisa não só ser concluído, como precisa ser concluído rapidamente. Khan et al [16] Destaca que não só o acesso veloz aos dados que estão por vir importam, como o armazenamento veloz destes dados não tratados para análises posteriores mais ágeis .
- **Veracidade e Valor:** A veracidade está relacionada com a qualidade dos dados obtidos enquanto o valor relaciona-se com o retorno do investimento [7]. Yin e Kaynak destacam que garantir um sistema que seja capaz de realizar o tratamento adequado dos dados, dando veracidade aos mesmos é de extrema importância pois quanto maior o volume de dados, mais difícil é filtrá-los e torná-los relevante [6]. Khan et al. dão muito destaque ao termo *valor* em relação a Big Data pois é nele que é mensurado de fato a qualidade dos dados obtidos levando-se em consideração todos os pontos anteriores, eles afirmam que quanto mais confiável for o retorno dos dados tratados, mais assegurado será o retorno financeiro ou intelectual sobre toda a estrutura montada para captação e tratamento desses dados [16].

É Importante frisar que todas estas variáveis se relacionam entre si, e configuram um ciclo, onde o sistema é retroalimentado, pois o produto, seja este de interesse financeiro ou intelectual, de alguma forma volta ao mercado específico. É possível entender melhor o ciclo através da **Figura 4:**

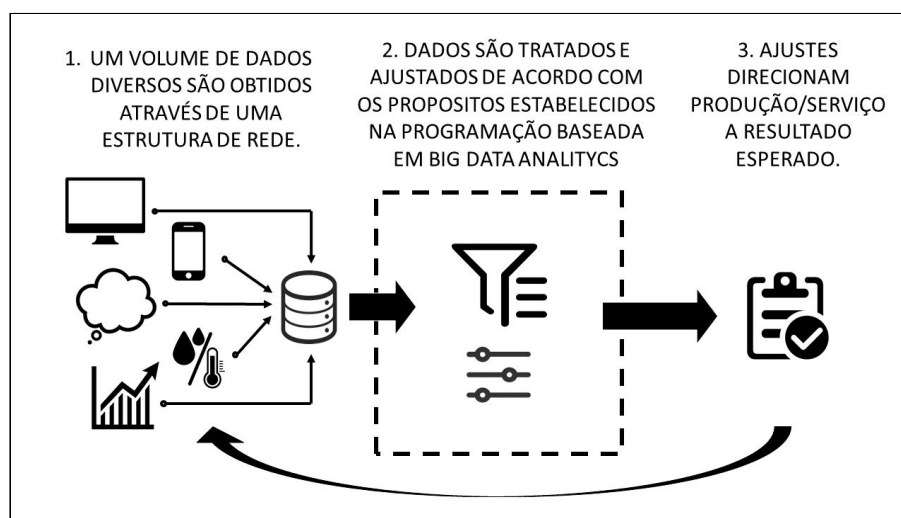


Figura 4. Uso da Big Data para tratamento de dados. **Autoria Própria (2019).**

3.2.1 INFRAESTRUTURA E CUSTOS PARA APLICAÇÃO DE MÉTODOS QUE USAM BIG DATA

Sabendo então o que é Big Data em um nível teórico, é necessário também avaliar o que o uso da mesma exige em questão de recursos e conhecimento específico para ser operacional. Barnes [20] especifica em seu artigo que “o processo de extrair, transformar e carregar dados em um banco de dados relacional é lento e requer que os dados sejam estruturados”. Ele cita Google e Facebook como exemplos quanto a dimensão do que seria a necessidade de informações em tempo real, os sistemas designados a suprir um rápido resultado na busca destas duas plataformas tem sido um imenso desafio. Para trabalhar com esses dados, métodos estatísticos são de fundamental importância, muitos destes fundamentados em KDD (*Knowledge Discovery in Databases*), sigla para descoberta de conhecimento em bases de dados [25].

Galvão e Marin [25] ainda especificam que o processo de KDD utiliza conceitos de base de dados, métodos

estatísticos, ferramentas de visualização e técnicas de *machine learning*, que dividem-se quanto aos dados usados nas etapas de seleção, pré-processamento, transformação, mineração de dados e interpretação destes. A *Machine Learning* traz força a esses processos por se adequar a demanda crescente de dados a se trabalhar que as corporações têm presenciado, concentrando o capital humano em tarefas ainda mais complexas.

Dentro do processo de KDD, temos a mineração de dados, as etapas anteriores a este, de seleção, processamento e transformação, muito se relacionam com a extração de conteúdo útil dos dados armazenados. A mineração de dados, como aponta Steiner et al [26], difere-se dos processos anteriores ao utilizar os dados para criação de modelos a partir das informações extraídas.

Steiner et al [26] afirma que a mineração dos dados ainda faz parte do processo KDD, porém, sendo a parte mais importante do mesmo, pois é através do desenho de padrões obtidos nessa etapa do método que as informações passam a ter muito mais valor e lógica para corporações e demais setores que o utilizem. A **Figura 5** ilustra o processo de KDD.



Figura 5. Etapas do processo KDD. Adaptado de [26] (2006).

Além disso, observa-se que a Big Data motivou a criação e a ampla disseminação de infraestruturas de software robustas para processamento e armazenamento desse grande volume de dados. Silva [27] cita o framework de código aberto *Apache Hadoop*, criado por Doug Cutting, como uma das ferramentas utilizadas nesse sentido. Esse framework tem a proposta de ser robusto e escalável quanto a armazenamento e processamento de dados, além de oferecer outras características como tolerância a falhas, transparência de distribuição, confiabilidade etc.

Silva [27] aponta como ponto forte da ferramenta a escalabilidade e segurança computacional. Seus dois principais serviços são o Sistema de Arquivos Distribuídos *Hadoop* e o *MapReduce*. Esse framework utiliza de alguns princípios abordados no KDD como seleção, pré-processamento e transformação dos dados a fim de entregar para quem o utiliza um resultado satisfatório de retorno. Há outros frameworks que utilizam destes princípios, como o *Spark*¹ e *Cassandra*².

Os dados precisam fazer sentido e ter uma designação do que de fato são para trazerem um resultado, ao qual Ling [21] se refere como “*Type*” (Tipo). No caso, precisamos compreender que o servidor é mero instrumento da tecnologia, ele armazena dados que serão requisitados seguindo uma lógica de linguagem específica. Ao sentirmos fome, não vamos diretamente ao supermercado, verificamos primeiro se há algo próximo que possa ser consumido. As linguagens combinadas ao servidor buscam essa “*Type*”, o resultado do que se busca pelo menor caminho possível, através da combinação de variáveis. Os métodos para requisição de dados em um banco de dados são chamados de “*queries*”, que são técnicas para requerer de forma estruturada o que está

1. Documentação do Apache Spark: <https://spark.apache.org/docs/latest/>
2. Documentação do Apache Cassandra: <https://cassandra.apache.org/doc/latest/>

armazenado em um banco de dados a fim de retornar um resultado lógico e com sentido [23].

Muito além dos servidores, a atividade computacional específica quanto a Big Data exige que os dados sejam processados antes de direcioná-los a esse banco de dados [16]. Ling especifica que as CPUs (Unidades Centrais de Processamento) convencionais por si podem não ser o suficiente para lidar com o paralelismo de dados que a Big Data exige, ao invés disso, um conjunto CPU e GPU (*Graphics Processing Unit*, em português unidade de processamento gráfico) deve ser mais adequado para tal tarefa, a CPU para as atividades sequenciais e a GPU para o tratamento de dados em múltiplos níveis [21]. A justificativa para o uso desse conjunto específico que sairia mais caro financeiramente que o uso de uma simples CPU é que as CPUs costumam ter apenas alguns núcleos de processamento, enquanto as GPUs possuem milhares, a capacidade computacional das GPU's entretanto, considerando a análise da Big Data, paga esse custo, conforme ilustrado na **Figura 6**.

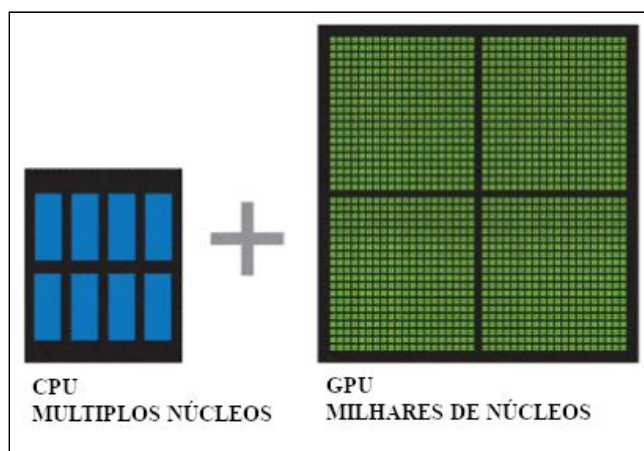


Figura 6. Comparação núcleos CPU e GPU. Adaptado de maketecheasier (2017).

Assim como Barnes [20] apontou a importância de sistemas robustos de armazenamento, Ling [21] vai além, não só robusto o sistema precisa também ser rápido para lidar com a escrita e leitura constante dos dados armazenados. O PCM (*Phase Change Memory*, em português Memória de Mudança de Fase) deve ser a solução para isso oferecendo velocidades superiores à de um disco rígido ou SSD convencional, mas esbarra na mesma característica que o conjunto CPUs e GPUs dedicadas, um maior custo em detrimento do uso de centrais de dados convencionais.

O que pode ser o maior entrave no Brasil e outros mercados emergentes quanto ao uso das tecnologias relacionadas ao Big Data, mais do que o custo em relação a aquisição de equipamentos robustos para uso local, é como será feito o transporte desses dados. Uma infraestrutura de rede robusta e dedicada para tal uso é o principal ponto em relação às tecnologias baseadas em Big Data, sem a comunicação destes com sistemas robustos o aprendizado de máquina pode não se desenvolver como esperado. Em artigo de Maia [22] sobre o uso de Big Data na agricultura brasileira, o mesmo aborda que em comparação aos Estados Unidos e Alemanha, maiores nomes no segmento, o Brasil esbarra na sua infraestrutura deficitária de transmissão de dados, pois os dados poderiam ser produzidos no local, mas as análises destes dados muitas vezes são feitas por empresas localizadas nos grandes centros urbanos.

3.3 DESAFIOS NA IMPLEMENTAÇÃO E IMPACTOS NO USO DE MÉTODOS BASEADOS EM BIG DATA.

Um dos desafios apresentados para uma implementação dos métodos relacionados ao tratamento de dados na indústria 4.0 é a necessidade que certos princípios sejam seguidos para garantir a precisão e eficiência do processo.

São oito os princípios que devem ser seguidos pela indústria, como aponta Yang [5]:

- Acessibilidade: Todos os envolvidos no projeto precisam ter as mesmas oportunidades de acesso aos padrões de tratamento de dados.
- Multilinguismo: Se tratando de um grande conjunto de dados, estes devem ter características muito próprias, como por exemplo uma grande variedade de idiomas. As tecnologias utilizadas no tratamento destes dados precisam ser capazes de lidar com esses dados.
- Segurança: Se tratando de dados, naturalmente que políticas que abordem a segurança destes precisam

ser seguidas a fim de garantir a total privacidade do mesmo desde sua origem.

- Privacidade: Seguindo o mesmo princípio da segurança de dados, políticas que garantam a segurança de quem fornece os dados são igualmente importantes.
- Subsidiariedade: A governança e os mais diretamente beneficiados pelas tecnologias de tratamento de dados devem estar o mais próximo possível, a fim de garantir transparência no processo.
- Uso de padrões abertos: Visando efetiva segurança no tratamento desses dados, o uso de padrões abertos, onde a base das soluções utilizadas pode ser analisada e implementada, é recomendável.
- Uso de software aberto: O software aberto, acessível para todos e colaborativo, é ferramenta de extrema utilidade para desenvolvimento de novas soluções para tratamento de dados que tendem a ficar cada vez mais complexos. Isso porque o uso tanto de soluções abertas a nível de padrões e softwares resultam em soluções multilaterais.
- Soluções multilaterais: Se tratando de problemas complexos, as soluções precisam vir de uma variedade de profissionais para que seja garantida a segurança e robustez nos sistemas de tratamento de dados.

Como já especificado por outros autores a respeito de Big Data [6, 7, 16], a definição da mesma dá-se por características ligadas diretamente a volume e receitação de dados variados, Yang [5] agora, ainda partindo dessas características, define limites e padrões que são necessários para atender a uma implementação de métodos que usem esses dados em vias domésticas, empresariais ou industriais. Definir desde o projeto inicial os padrões a serem utilizados permite que sejam calculadas as medidas a serem tomadas para garantir a segurança dos dados, que não são, por vezes, unicamente propriedade daqueles que o interceptam, como também mensurar o custo financeiro envolvido para que seja de fato implementado um sistema dessa magnitude, qual estrutura escolher e quais contratos com terceiros firmar? O respeito com os dados obtidos e a transparência que eles são utilizados e o porquê são utilizados é medida que precisa ser debatida e preferencialmente adotada por aqueles que usem de dados que não sejam estritamente públicos de seus detentores ou que assim eles os considere.

3.3.1 BIG DATA NO BRASIL

Sobre a Big Data no Brasil, de acordo com a Confederação Nacional das Indústrias (CNI), em 2016, 63% das grandes empresas, sejam elas indústrias ou empresas voltadas ao comércio e mercado de serviços, já usavam tecnologia digital visando produtividade. Em 2018 essa mesma utilização subiu para 73%, ainda com tecnologias voltadas para produção [11]. Como o Brasil ainda é um país com baixos níveis de automação, é comum que as empresas se voltem primeiramente ao operacional antes de seguir nas demais etapas, contudo, tal característica pode ser extremamente danosa à competitividade da indústria e comércio no geral se considerarmos conforme Coelho [2] que em 2020, já teremos no mundo um nível de sofisticação relacionada a big data muito grande. A CNI também aponta no mesmo relatório que ao questionar empresas que pretendem investir em tecnologia digital, 79% informa que possui foco na produção ou gestão do negócio, onde o foco em inovação, área onde métodos baseados em Big Data apresentam grandes possibilidades, ainda é pouco explorado.

Um dos casos estudados com forte relação quanto ao processo de renovação da indústria apontado pela CNI tratava da indústria 4.0 na gestão da demanda hídrica de um frigorífico [13]. Foi identificado na pesquisa que a estrutura na cadeia produtiva carecia de elementos que fornecessem maior controle sobre o consumo de água por setor, onde o consumo era medido por área. Tal situação vai em direção contrária às tendências de maior investimento na área de qualidade de dados [3], a inexistência dos dados para cada setor de produção gerava um uso ineficiente de água no local. A recomendação a priori foi relacionada a instalação de sensores e hidrômetros, que naturalmente controlariam a saída de água e mediriam o volume de água envolvido em cada processo.

Como o interesse era diretamente o controle de vazão e o armazenamento de dados, estes sistemas ciber-físicos, que utilizam de uma interface digital para controle conforme apontado por Coelho [2], poderiam trazer à tona dados históricos de consumo que antes não seriam apontados no processo.

3.3.2 IMPORTÂNCIA DA BIG DATA PARA INDÚSTRIA E GOVERNO

Schmidt et al. [14] aponta que muitas companhias encaram o desafio que é evoluir e entender a diversidade de conceitos que são relacionados à indústria 4.0 e com estes desenvolver estratégias de negócio próprias. Ele aponta ainda um dilema que é enfrentado pelas organizações: esperar demais para iniciar sua implementação na indústria 4.0 ou começar rapidamente? É de extrema importância delimitar o que é urgente na implementação de um esquema de produção baseado na indústria 4.0. Sendo essencial que os elementos, diferente de uma linha de produção que a antecede, interajam entre si. A tecnologia Big Data será determinante nesse sentido, considerando que o objetivo na implementação de dispositivos que interagem entre si, ou indústrias inteligentes, depende da movimentação de uma grande rede de dados entre sistemas [6]. Sistemas estes que não só sejam

capazes de tratar os dados, mas de os interpretar e agir sobre eles a fim de alcançar um objetivo de produção ou prestação de serviço.

Há também tecnologias chamadas de governo móvel (m-government) [9], estas se dão por um esforço conjunto da iniciativa público-privada para universalizar a comunicação de ocorrências mesmo por um público mais marginalizado ou com dificuldade de locomoção. A análise combinatória das informações, transmitidas desde aplicativos e sites próprios quanto de terceiros, como redes-sociais de largo alcance, possibilita uma gestão inteligente dos recursos permitindo que este seja direcionado para as demandas mais urgentes, economizando dinheiro público e aumentando a satisfação da sociedade como um todo. Estas tecnologias enfrentam maior impacto na conectividade da população, que é o maior instrumento na atualização destes banco de dados.

Além do robusto banco de dados, frisa-se também na literatura estudada que a integração com sistemas ciber-físicos orientados a internet das coisas é essencial para o pleno funcionamento de uma cadeia de produção baseada em indústria 4.0 [15]. Os sistemas ciber-físicos, aliados a uma amplitude de equipamentos conectados a uma rede que tem seus dados armazenados no banco de dados empresarial ou aberto, poderiam determinar, sem a intervenção humana, por exemplo, o grau de satisfação de clientes através de um histórico de produção, e por este, atualizar parâmetros visando otimizar o produto final pelo menor custo possível. O contato com o público alvo ou os responsáveis pela produção, em um cenário em que a IoT é plenamente aplicada, poderia se dar através de páginas web, softwares ou aplicativos orientados a este consumidor ou responsável, como apontado por Lemos e Araújo [6], a respeito do uso de tecnologia de comunicação móvel para envio de dados a órgãos públicos e empresas privadas.

3.3.3 DESAFIOS SOCIAIS E ÉTICOS NO USO DA BIG DATA

Cuquet [10] aponta que a aplicação de métodos baseados em Big Data possui vasto campo de implicações e problemas, sejam estes sociais, éticos ou legais. O mesmo relaciona esses problemas em especial a segurança e privacidade dos dados em um mundo conectado e em que a transferência de dados em uma esfera global ocorre. A garantia de segurança e uso desses dados mais do que simplesmente assegurada, precisa ser debatida nas organizações, a fim de garantir que sejam supridas exigências das legislações de cada país quanto a privacidade e segurança dos dados.

Zheng aponta ainda que outro desafio na implementação de métodos ligados a Big Data em companhias atualmente se dá pela relutância ou desinteresse na melhoria do registro de dados internos da própria organização, partindo de opiniões que sugerem que tal atividade não agrega valor ao objetivo final e sendo apenas uma etapa do processo produtivo, entretanto, estas medidas devem mudar à medida que uma indústria que possui como base a customização em massa comece a surgir [15].

3.3.4 BIG DATA NA VISÃO DA INDÚSTRIA

A visão da indústria tende a voltar seus olhos entretanto para o chamado aprendizado de máquina ou *machine learning*, e é nesse nível que os dados internos armazenados pela organização e dados externos serão valorizados [17]. Outro desafio na implementação de estruturas que usem da Big Data relaciona-se diretamente ao capital humano, Groger [18] afirma que o mercado apresenta grande número de profissionais com conhecimento sobre o negócio ou sobre a análise avançada de dados em separado, mas o que o mercado precisa para o futuro é de profissionais com conhecimento industrial e estatístico para aplicação de métodos relacionados a análise de dados no campo empresarial. Mashelkar [19] concorda com a visão de Groger. O mesmo aponta que com a evolução tecnológica que ocorre há várias décadas, é normal que novas modalidades de trabalho surjam e outras se tornem obsoletas, o mesmo afirma que “empregos desaparecem, trabalho não”.

3.4. PERSPECTIVAS FUTURAS PARA O USO DA BIG DATA

Diante das informações analisadas durante a pesquisa, é possível verificar de acordo com o material acadêmico publicado que uma estrutura adequada de análise de dados voltada ao processo produtivo ou de distribuição de serviços não só é necessária como indispensável quando se trata de indústria 4.0, sendo parte fundamental dos pilares e necessário para que faça sentido o termo indústria inteligente, sendo essa totalmente conectada e adotando a customização em massa na produção de bens e serviços com mínima intervenção humana. Após a leitura dos artigos, em todos os que foram publicados e avaliados após o campo metodologia, exceto de Schneider et al. [13], O termo Big Data é citado com destaque, por vezes com sessão exclusiva ou já tratando no título de sua importância.

As possibilidades de uso da tecnologia são tão extensas que a aplicação de métodos de análise baseados em Big Data já tem sido discutida em pequenas comunidades conectadas, ligando-se diretamente ao conceito de cidades inteligentes, buscando o controle e sustentabilidade de todo o “ecossistema” das mesmas através de ferramentas tecnológicas, de simples smartphones à câmeras e sensores mais avançados [8]. A presença de tais instrumentos permitiria por exemplo um maior controle e agilidade na solução de problemas de mobilidade ou estrutura na comunidade em que tais serviços e tecnologias estão inseridos. A presença de aplicativos que resultem em comunicação direta com os órgãos responsáveis por tais demandas são exemplo disso. Na cidade de Mossoró-RN, como podemos ver na **Figura 7**, a prefeitura já usa essa possibilidade por meio de um aplicativo onde a população pode solicitar serviços e ao mesmo tempo gera um banco de dados sobre tais oportunidades. Mesmo que ainda não seja aplicado o uso de uma inteligência artificial robusta no sistema, a criação de um banco de dados traz um horizonte possível.



Figura 7. Aplicativo “Mossoró Conectada” permite uma gestão mais eficiente da cidade e gera um banco de dados da região. **Fonte:** CGS Consultoria em Tecnologia (2019).

É interessante que o único artigo voltado para Indústria 4.0 que não tenha abordado Big Data tenha aplicação em uma indústria brasileira, indo de encontro aos dados da CNI [11] que apontam que por mais que cresça o interesse do mercado brasileiro em relação a qualidade de dados, a estrutura industrial no Brasil ainda tem alguns pontos urgentes a tratar, como a automação nas linhas de produção e o desenvolvimento de uma estrutura de servidores para tratamento de dados que talvez ainda não seja realidade para aplicação em boa parte do país.

Sobre a importância vital da Big Data para a indústria 4.0, é importante salientar ainda, que a implementação de métodos que tratam de dados relaciona-se diretamente com a estrutura que um projeto de indústria 4.0 objetiva, onde os elementos agem de forma independente, permitindo assim uma produção diversificada, porém que interliga-se por um poderoso sistema computacional e que controla as ferramentas produtivas visando atingir um objetivo final, objetivo este que pode ser alterado de acordo com os dados obtidos e tratados no início do ciclo, como mostrado na seção 2.2.

A maior problemática que a aplicação de métodos ligados a Big Data apresenta para o futuro é o desinteresse de companhias do mundo todo na importância que a qualidade de dados apresenta, como apontado pelo executivo David Wang [17], usando-os por vezes, mesmo em um ambiente altamente automatizado, como ferramenta de análise para a produção e não para nortear os fundamentos da corporação, ignorando oportunidades que são apresentadas diariamente pelo vasto e variado mar de dados que está disponível desde a explosão do uso de eletrônicos e popularização de redes sem fio de comunicação vindas neste século.

Além dos impactos tecnológicos, os impactos sociais e legais da aplicação de tais métodos baseados em análise de grandes volumes de dados, de diversas fontes, como redes sociais, faz com que um padrão rígido e eficiente sobre o tráfego e uso desses dados precisa ser debatido nas organizações. O surgimento de profissionais com conhecimento de mercado, mas também estatístico e voltado a lógica de programação deve crescer nas próximas décadas conforme apontado por Groger [18], pois será tendência para o futuro e naturalmente, como em toda revolução, traz consigo novos postos de trabalho e profundas mudanças na estrutura econômica vigente, pois os dados são direcionados a massa, a produção de riqueza por si só não faz sentido [19].

Outro ponto que pode dificultar a implementação efetiva de métodos baseados em Big Data nas organizações trata-se do *status quo* de executivos quanto ao *modus operandi* da empresa ou indústria, além da desconfiança do que é gerado por uma inteligência artificial, executivos podem rejeitar os dados por práticas de mercado voltadas a experiência e *feeling* do profissional. Galbraith cita o caso Dell como exemplo no surgimento da bolha da internet no início do século XXI. A companhia adotou o método de vendas online imediatamente, vendo o mesmo como oportunidade ao seu modelo de negócios enquanto a HP, companhia do mesmo ramo, rejeitou completamente o modelo perdendo mercado posteriormente [24], a rejeição a tecnologia por mais que estranha para quem esteja introduzido a mesma é comum no meio empresarial, principalmente levando em consideração o quão burocrática a organização possa ser.

A existência de profissionais experientes e com *feeling* de mercado, naturalmente não deixa de ser importante desde que não resulte no bloqueio das ferramentas baseadas na tecnologia, atuando em sinergia. Como apontado por Fagundes, o volume de dados para nossa época é exponencial e muito além da capacidade humana, ou mesmo de uma larga equipe de indivíduos, de ser analisada [6]. É preciso compreender que os dados internos de uma organização, armazenados de forma individual e catalogada, ou meramente vinculados a um software específico da mesma para comunicação, não mais são o suficiente para direcionar a produção, dessa forma, uma análise lógica de dados que venha de fontes diversas, desde redes sociais a dados históricos de produção da companhia em situações diversas se faz necessário.

A tendência apontada por Groger [18] que atualmente há poucos profissionais que de fato sejam aptos a lidar com a tecnologia Big Data em sua plenitude no mercado de trabalho é extremamente importante, pois a tecnologia Big Data não é meramente analítica e deve haver um estudo e se ter certa habilidade no assunto estudado para garantir que o sistema de fato oferece resultados coesos através dos bancos de dados e métodos de requisição ou “*queries*” anteriormente abordados [21]. Tais pontos elencados aqui são abordados de forma mais direta e simples na **Tabela 1**, de autoria própria:

Tabela 1. Desafios a enfrentar para plena aplicação da Big Data. (Autoria Própria).

RESUMO DOS DADOS LEVANTADOS	
Quais os impactos identificados quanto a utilização da Big Data nas organizações?	Maior customização de produtos e serviços voltados ao público e rápida evolução na área de inovação, maior custo organizacional em relação a segurança, tratamento e armazenamento de dados, surgimento de debate a respeito de privacidade e segurança de dados, reformulação dos padrões hierárquicos das organizações e surgimento de novas profissões voltadas a ciência de dados.
Tecnologias necessárias para o uso da Big Data.	Servidores, locais ou contratados em nuvem para armazenamento dos dados, uso de métodos estatísticos para o tratamento destes dados seguindo conceitos como o KDD, aplicações de fácil acesso a nível homem para que os dados tratados sejam compreendidos e uso de ferramentas diversas para alimentação do sistema de dados, em um sistema cyber-físico bem fundamentado.
Maiores desafios a ser superados para possibilitar uma implementação efetiva	Infraestrutura de rede, capital humano com conhecimento técnico, alto custo para aquisição de equipamentos (hardware) específicos.
Setores beneficiados	Todos os setores tendem a beneficiar-se do uso da Big Data em um futuro próximo, trazendo consigo novas profissões e demandas quanto ao capital humano.

4. CONCLUSÕES

Ao fim do estudo e considerando todas as especificidades apresentadas no material abordado, é possível afirmar quanto a desafios que, em especial nos países emergentes há ainda muito o que ser feito para plena utilização de métodos baseados em Big Data e consequentemente para plena operação da Indústria 4.0 nesses países. Cientistas de dados tendem a crescer no mercado e aqueles que além do viés analítico possuir também conhecimento empírico a mercados específicos devem se destacar.

O investimento em infraestrutura de comunicação também não pode ser ignorado, devendo ser bilateral e priorizado não só pela indústria específica como pelo governo local para potencializar a competitividade da mesma ante outros mercados. O Brasil em específico, nos artigos que abordam a implementação de estrutura de dados como o da gestão hídrica do frigorífico [13] e o uso de soluções em Big Data no campo [22], tem como maior deficiência sua rede de comunicação com e sem fio. E esse problema em especial não pode ser resolvido de forma eficiente e ágil somente através da iniciativa privada, em especial quando trazemos a problemática para negócios pequenos e em locais de difícil acesso, que possuem um maior potencial de crescimento e que precisam dessa estrutura para enviar os dados coletados para um datacenter mais robusto e capaz de tratá-los e prever os resultados obtidos, igualmente retornando de forma ágil para o produtor ou equipe industrial.

Sendo apontado por todos os artigos que abordaram o método como fundamental para funcionamento do mesmo, o uso de dados por si não é o suficiente e sim a transmissão dos mesmos para tratamento e análise. O uso de estruturas robustas a depender do tamanho de dados produzidos também é um empecilho para pleno funcionamento da Big Data em especial nos mercados emergentes pois é diretamente relacionado ao uso de hardware de ponta como GPUs e PCMs anteriormente abordados.

A **Tabela 2** mostra resumo dos principais desafios envolvidos no uso de Big Data nas organizações. Ela foi criada a partir do levantamento feito no desenvolvimento deste trabalho:

Tabela 2. Considerações finais sobre desafios para uso da Big Data de acordo com artigos avaliados. (Autoria própria).

DESAFIOS NA APLICAÇÃO DA TECNOLOGIA BIG DATA	
ASPECTO	DESAFIO
CAPTURA DE DADOS	Investimento na estrutura organizacional e em tecnologias de análise e tratamento de dados [25]. É necessário garantir a veracidade dos dados antes de usá-los como referência em uma indústria/comércio que ainda os cataloga de forma linear.
INFRAESTRUTURA	Em especial em países em desenvolvimento, a estrutura deficitária de redes físicas e por satélite para comunicação de dados é o maior desafio, pois envolve alto custo financeiro tanto por parte das empresas privadas quanto do governo local para de fato ser efetivo. O investimento em tecnologia a nível de hardware in loco pode ser necessário a depender do volume de dados.
CAPITAL HUMANO	A popularização da ciência de dados a nível comercial e para indústrias de pequeno porte devem trazer uma demanda ainda inexistente de profissionais devidamente qualificados para atividades voltadas ao uso da tecnologia big data. Esses profissionais além de conhecimento técnico preferencialmente devem conhecer aspectos administrativos e organizacionais ou atuar em sinergia com profissionais dessa área.
SEGURANÇA DE DADOS	Com o uso de dados de fontes diversas graças a popularização também da internet das coisas, políticas públicas de privacidade e segurança de dados devem surgir, portanto, adaptar-se a padrões de segurança e respeito a privacidade de usuários deve ser algo que as organizações devem se antecipar na implementação da tecnologia.

Esse trabalho serve como base para aqueles que pretendam não só entender o que é Big Data como também compreender sua importância e ferramentas na implementação da chamada indústria 4.0.

Devido a amplitude que o assunto engloba e do tempo disponível para realização do trabalho, nem toda a literatura foi abordada. Porém, as informações levantadas são de relevância para introduzir o tema Big Data, sendo um instrumento de aperfeiçoamento pessoal e profissional que indica tendências e impactos quanto ao assunto.

Como sugestão para projetos futuros, o desenvolvimento de um método de análise e estudo em local que tais ferramentas já estejam operando ou um novo estudo abordando de forma mais ampla a literatura são possibilidades.

5. REFERÊNCIAS BIBLIOGRÁFICAS

1. CAI, Li; ZHU, Yangyong. **The Challenges of Data Quality and Data Quality Assessment in the Big Data Era**. Data Science Journal 14, Maio de 2015. Disponível em: <https://www.researchgate.net/publication/277943983_The_Challenges_of_Data_Quality_and_Data_Quality_Assessment_in_the_Big_Data_Era>. Acesso em: 01 de novembro de 2018.
2. COELHO, Pedro. **Rumo à Indústria 4.0, Universidade de Coimbra**. Disponível em: <https://estudogeral.uc.pt/bitstream/10316/36992/1/Tese_Pedro_Coelho_Rumo_à_Industria_4.0.pdf>. Acesso em: 12 de dezembro de 2018.
3. CAMPOS, F.; SILVA, I.; **Big data e cloud computing, perspectivas de uso em engenharia de produção: uma análise bibliométrica de 2004 – 2014**. ENEGEP, 2015. Disponível em: <http://www.abepro.org.br/biblioteca/TN_STO_213_261_27525.pdf>. Acesso em: 01 de novembro de 2018.
4. LORENZ et al; **Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries**. BCG, 2015. Disponível em: <https://www.bcg.com/pt-br/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx>. Acesso em 22 de janeiro de 2019.
5. LU, Yang; **Industry 4.0: A Survey on Technologies, Applications and Open Research Issues**. Journal of Industrial Information Integration, 2017. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2452414X17300043>>. Acesso em: 22 de janeiro de 2019.
6. YIN, Shen; KAYNAK, Okyay. **Big Data for Modern Industry Challenges and Trends**. Proceedings of the IEEE, vol. 103, Fevereiro de 2015. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7067026>>. Acesso em: 24 de fevereiro de 2019.
7. FAGUNDES, et al. **A produção científica sobre qualidade de dados em big data: um estudo na base de dados web of Science**; Revista digital de biblioteconomia e ciência da computação. p. 3-6. 2017. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8650412/pdf>>. Acesso em: 24 de outubro de 2018.
8. SUN, et al. **Internet of Things and Big Data Analytics for Smart and Connected Communities**. IEEE Access, Vol. 4, p. 766-773, 2016. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7406686>>. Acesso em: 01 de novembro de 2018.
9. LEMOS, André; ARAÚJO, Nayra Veras. **Cidadão Sensor e Cidade Inteligente: Análise dos Aplicativos Móveis da Bahia**. Revista Famecos, Porto Alegre, v. 25, n. 3, p. 1-19, setembro, outubro, novembro e dezembro de 2018. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/revistafamecos/article/download/28708/17204>>. Acesso em 12 de dezembro de 2018.
10. CUQUET, et al. **Societal impacts of big data: challenges and opportunities in Europe**. Cornell University, 2017. Disponível em: <<https://arxiv.org/abs/1704.03361>>. Acesso em: 20 de dezembro de 2018.

11. CONFEDERAÇÃO NACIONAL DA INDÚSTRIA. **Investimentos em Indústria 4.0**, CNI, 2018. Disponível em: <https://bucket-gw-cni-static-cms-si.s3.amazonaws.com/media/filer_public/8b/0f/8b0f5599-9794-4b66-ac83-e84a4d118af9/investimentos_em_industria_40_junho2018.pdf> Acesso em: 18 de fevereiro de 2019.
12. SILVA, Cassandra Ribeiro de O. **Metodologia e Organização do Projeto de Pesquisa**. Guia prático. Fortaleza, 2004 Disponível em: <<http://joinville.ifsc.edu.br/~deborapac/Metodologia%20e%20Organiza%C3%A7%C3%A3o%20do%20Projeto%20de%20Pesquisa%20CEFET%20CE.pdf>>. Acesso em: 12 de dezembro de 2018.
13. SCHNEIDER et al., **a aplicação do conceito, restrito, de indústria 4.0 à gestão da demanda hídrica em um frigorífico**. ENEGEP, 2017. Disponível em: <http://www.abepro.org.br/biblioteca/TN_STO_246_423_32903.pdf>. Acesso em: 18 de fevereiro de 2019.
14. SCHMIDT, Rainer et al. **Industry 4.0 - Potentials for Creating Smart Products: Empirical Research Results**, Conference Paper · June 2015. Disponível em: <https://www.researchgate.net/publication/274894802_Industry_40_Potentials_for_Creating_Smart_Products_Empirical_Research_Results>. Acesso em: 24 de fevereiro de 2019.
15. ZHENG, Pai et al. **Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives**. Frontiers of Mechanical Engineering. Junho de 2018, Volume 13, p.137–150. Disponível em: <<https://link.springer.com/article/10.1007/s11465-018-0499-5>>. Acesso em: 25 de fevereiro de 2019.
16. KHAN, Ali-ud; UDDIN, Muhammad; GUPTA, Navarun. **Seven V's of Big Data Understanding Big Data to extract Value**. IEEE, Conference of the American Society for Engineering Education, 2014. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/6820689>>. Acesso em: 01 de março de 2019.
17. WANG, David. **Building Value in a World of Technological Change: Data Analytics and Industry 4.0**. IEEE DOI 10.1109/EMR.2018.2809915. Disponível em: < <https://ieeexplore.ieee.org/document/8334872> >. Acesso em: 01 de março de 2019.
18. GROGER, Cristoph. **Building an Industry 4.0 Analytics Platform**, Approaches and Future Research Directions. Datenbank-Spektrum 2018. Springer, Berlin Heidelberg. Disponível em: < <https://doi.org/10.1007/s13222-018-0273-1> >. Acesso em 02 de março de 2019.
19. MASHELKAR, R. A. **Exponential Technology, Industry 4.0 and Future of Jobs in India**. Review of Market Integration 10ª ed, 2018. Disponível em: < <http://journals.sagepub.com> >. Acesso em: 02 de março de 2019.
20. BARNES, Dan. **Report: Data analytics - Computing infrastructure - Big data becomes big business**. The Banker, novembro de 2012. Disponível em: <<http://link.galegroup.com/apps/doc/A310969467/AONE?u=capes&sid=AONE&xid=0f5fc2c3>> . Acesso em: 02 de abril de 2019
21. LING, Liu. **Computing infrastructure for big data processing**. Georgia Institute of Technology, Atlanta, 2013. Disponível em: <<https://link.springer-com.ez13.periodicos.capes.gov.br/article/10.1007/s11704-013-3900-x>>. Acesso em: 03 de abril de 2019.
22. MAIA, Humberto. **Da terra brotam os dados: o uso do big data no campo promete aumentar a produtividade. A tecnologia está chegando ao Brasil, mas esbarra nos problemas que as propriedades enfrentam da porteira para fora**. Portal Exame, 2014. Disponível em: <<http://link.galegroup.com/apps/doc/A394684202/AONE?u=capes&sid=AONE&xid=c6abb114>> Acesso em: 06 de abril de 2019.
23. DEL ROBERT, Rossi; SPRAGENS, John. **Exploring the data maze: query tools offer a variety of ways to extract answers from your SQL database**. InfoWorld, novembro de 1993, pág. 75+. Disponível em:

<<http://link.galegroup.com/apps/doc/A14623021/AONE?u=capes&sid=AONE&xid=82de6d20>>. Acesso em: 07 de abril de 2019.

24. GALBRAITH, Jay. **Organization design challenges resulting from big data**, JOD, 13 de fevereiro de 2014. Disponível em: <<https://doaj.org/article/d294f54a2f1c44fa99c80073d4057ccf>>. Acesso em: 08 de junho de 2019.

25. GALVÃO, Noemi; MARIN, Heimar. **Técnica de mineração de dados: uma revisão da literatura**. Acta Paulista de Enfermagem, vol. 22, núm. 5, outubro, 2009, pp. 686-690 Escola Paulista de Enfermagem São Paulo, Brasil. Disponível em: <<https://www.redalyc.org/pdf/3070/307023846014.pdf>>. Acesso em: 17 de agosto de 2019.

26. STEINER et al. **Abordagem de um Problema Médico por meio do Processo de KDD com ênfase à Análise Exploratória dos Dados**, Scielo, Gestão e produção, v.13, n.2, p. 325-337, 2006. Disponível em: <<http://www.scielo.br/pdf/gp/v13n2/31177.pdf>>. Acesso em: 18 de agosto de 2019.

27. SILVA, Bruno. **Apache Hadoop**, Universidade Federal de Santa Catarina. Disponível em: <<https://dcomp.sor.ufscar.br/verdi/topicosCloud/Hadoop.pdf>>. Acesso em: 19 de agosto de 2019.



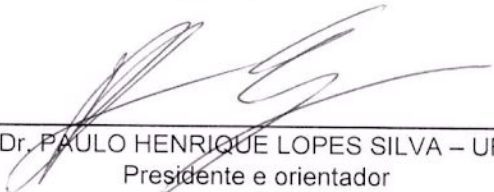
MINISTÉRIO DA EDUCAÇÃO
Universidade Federal Rural do Semi-Árido – UFERSA
Centro de Ciências Exatas e Naturais – CCEN

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

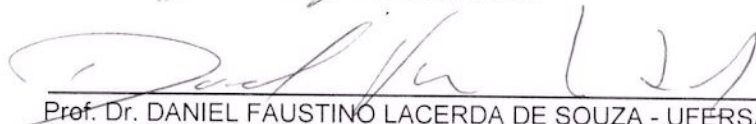
Às quatorze horas do dia 14 de agosto de 2019, no LAACOSTE – Departamento de Computação da UFERSA – Universidade Federal Rural do Semi-Árido, reuniu-se a Banca Examinadora de defesa de trabalho de conclusão de curso de autoria do aluno **JHONATAN KAYO SILVA SENA**, aluno do curso de Bacharelado em Ciência e Tecnologia desta universidade, N° de matrícula **2014010988**, com o título **“UM ESTUDO SOBRE A IMPORTÂNCIA DA TECNOLOGIA BIG DATA E PERSPECTIVAS FUTURAS DE SUA APLICAÇÃO NA INDÚSTRIA”**. A Banca Examinadora ficou assim constituída por três membros: Prof. Dr. PAULO HENRIQUE LOPES SILVA, presidente da banca e orientador do Trabalho de Conclusão de Curso; Prof. Dr. DANIEL FAUSTINO LACERDA DE SOUZA MEMBRO e Profa. Ma. WELLIANA BENEVIDES RAMALHO, como membros. Concluída a defesa, procedeu-se o julgamento pelos membros da banca examinadora, tendo o aluno obtido as seguintes notas: 90; 90; 90. Apuradas as notas verificou-se que o aluno foi aprovado com média geral 90. E para constar, eu, PAULO HENRIQUE LOPES SILVA, lavrei a presente ata que, após lida e aprovada pelos membros da banca examinadora, será assinada por todos.

Mossoró, 14 de agosto de 2019.

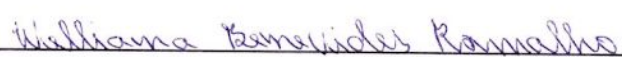
Assinatura dos membros da Banca Examinadora.



Prof. Dr. PAULO HENRIQUE LOPES SILVA – UFERSA
Presidente e orientador



Prof. Dr. DANIEL FAUSTINO LACERDA DE SOUZA - UFERSA
Primeiro Membro



Profa. Ma. WELLIANA BENEVIDES RAMALHO - UFERSA
Segundo Membro